



Digitaaldokumentide arhiveerimine ja ülekandmine arhiivi

Protage töötuba

Tallinn, 03.04.2008
Kuldar Aas, Raivo Ruusalepp



Päevakava

11.30 – Sissejuhatus

11.40 – Mis on Protage?

12.00 – Mis või kes on agendid?

12.15 – Digitaaldokumentide arhiveerimise
planeeritav töövoog Eestis ja selle võimalik
automatiseerimine

12.30 – 13.30 Arutelu

13.30 – 14.00 Kohvipaus

14.00 – 16.00 Arutelu jätkub

16.00 – 16.30 Kokkuvõte ja hüvastijätt



Eesmärk

- Esmane: saada tagasisidet Protage'i töö jätkamiseks
- Teisene: koguda arvamus Rahvusarhiivi digitaalarhiivi edasiseks arendamiseks

Mõningaid reegleid

- Miski pole võimatu – vaja on ideid mitte lahendusi!
- IT on liiga keeruline? – unusta kõik halb, mida arvutid Sulle teinud on
- Peamine küsimus – mis teeks Sinu töö ja elu lihtsamaks?

Dokumenteerimisest

- Arutelu lindistatakse
- Arutelu dokumenteeritakse kirjalikult
- Osalejatele eestikeelne kokkuvõte, selle kommenteerimine ja täiendamine kuni 09.04 on igati tervitatav!
- 11.04.2008 saadetakse ingliskeelne kokkuvõte projekti partneritele



Mis on Protage?



PROTAGE

*PR*eservation *O*rganizations *T*esting **AG**ent **E**nvironments

7th Framework Programme

Grant agreement no 216746





Euroopa Komisjoni teemad

- **ICT Call 1**
- **Objective 3.4.1.1 Digital libraries and technology-enhanced learning**
 - Medium term:
 - Large-scale European-wide digital libraries
 - Long term:
 - Radically new approaches to digital preservation

http://cordis.europa.eu/fp7/ict/telearn-digicult/digicult-call1_en.html



Protage projekt

- Project type: Small or medium-scale focused research project (STREP)
Start date: 01 November 2007
Duration: 36 months
EU funding: € 2 021 900
Number of partners: 7
Project coordinator: Riksarkivet (National Archives), Sweden



Protage konsortsium

- The National Archives, Sweden (RA)
- Luleå University of Technology (LTU)
- National Archives of Estonia (NAE)
- Fraunhofer Gesellschaft (FRAUNHOFER)
- University of Bradford (UNIBRAD)
- EASY Innova S.L. (EASY)
- Giunti Labs S.r.l. (GILABS)



Protage eesmärgid

- research about the potential of software agent ecosystems to support the automation of digital preservation tasks;
- demonstration of the technical feasibility of such a system by means of a prototype;
- analyse how this system can be implemented in various organisational environments;
- explore the possible integration of PROTAGE solutions with other or existing digital preservation environments;
- explore synergies with other RTD activities dealing with digital preservation.



Rahvusarhiiv Protage'is

- Aitab määratleda loodava prototüübi rolli ja ülesannet
- Osaleb prototüübi kasutuslugude väljatöötamises
- Vastutab prototüübi testimise eest (WP4) nii arhiivi kontekstis kui ka laiemas kasutajate ringis
- Osaleb projekti tutvustamises

Tänane arutelu

- Koguda lisa-teadmist ja vajadusi valdkonna praktikutelt
- Vajaduste, soovide, unelmate ja nõuete põhjal täiendatakse Protage lahenduse prototüübi lähteülesannet ja võimalusel arendatakse vajalikud funktsionaalsused
- Vorm: vaba diskussioon, arutelu struktuuri jaoks on võimalik järgida dokumentide arhiivi üleandmise töövoogu



Mis vői kes on agendid?

Mis on agent?

Agent on arvutisüsteem, mis on võimeline tegutsema iseseisvalt oma kasutaja või omaniku huvides. Selle asemel, et talle peaks ütlema kuidas konkreetses olukorras käituda, oskab agent ise olukorda hinnata ja õigeid otsuseid langetada.

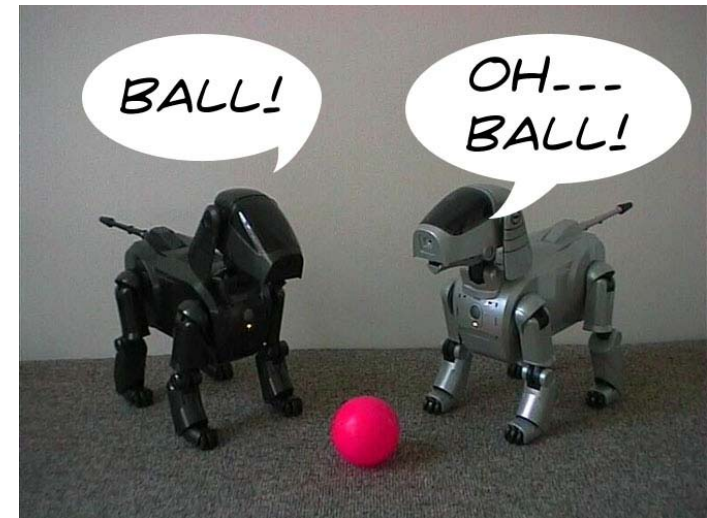


Olulised märksõnad

- Pidevus – agent töötab, teda ei käivitata
- Autonoomsus – tegevuste prioritiseerimine ja otsuste tegemine ilma inimese osaluseta
- Sotsiaalsus – suhtlemine teiste agentide ja süsteemidega, koostöös probleemide lahendamine
- Reaktiivsus – pidevalt võrreldakse tegevusi ümbritseva keskkonnaga ning selle muutumisel muudetakse tegevusloogikat

Mis on agenttehnoloogia?

- Agenttehnoloogia on tark- ja riistvaraliste agentide arendusmeetodite kogum
- Tarkvaralised agendid – agendi omadustega programmid
- Riistvaralised agendid – õppivad robotid

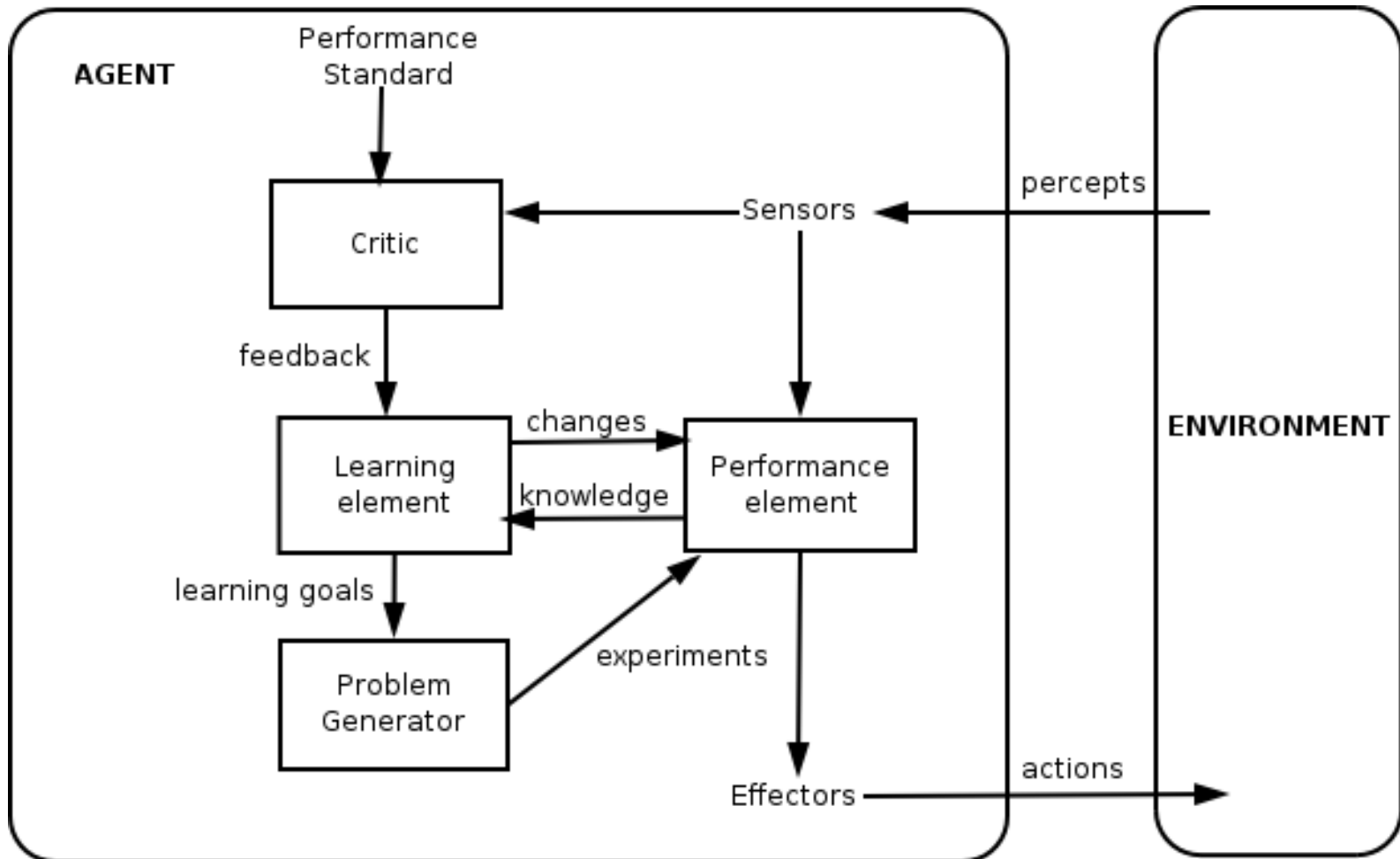




Agentide kasutamine

- Protage käsitleb ainult tarkvaralisi agente, seega robotitest täna ei räägi
- Tarkvaralisi agente:
 - Kasutatakse enamasti peidetuna teiste süsteemide sisse
 - Levinuimad agendid: ostusoovitused veebis
 - Aga kasutatakse ka liikluse kontrollimisel, tootedisainis, tootmisliinidel ...

Intelligentised agendid





Agentide koostöö

- Protage'i raames arendatakse omavahel suhtlevate agentide “ühiskond”
- Igal “ühiskonna” liikmel on oma spetsiifilised oskused
- Oluliseim osa usaldusel – millise “ühiskonna” liikme arvamust arvestada?
- Agent areneb ja õpib – mida kauem on ta olnud “ühiskonna” liige ja mida suurem on “ühiskond”, seda kvaliteetsemaid otsuseid ta teeb

Agendid ja arhiivindus

- Agentide digitaalarhiivides rakendamise võimalikkuse uurimisega on hakatud tegelema viimasel paaril aastal, kasutuses veel ühtegi lahendust ei ole
- Teoreetiline võimalus – agente on mõttekas kasutada igas tegevuses, mille puhul on tähtsam küsimus “*Kuidas?*” mitte “*Miks?*”



Töövoog digitaaldokumentide arhiveerimisel





Üldstruktuur sama

- Taotluse ettevalmistamine
- Taotluse esitamine arhiiviasutusele
- Vajadusel hindamine
- Korrastamine ja kirjeldamine
- Nimistu kooskõlastamine
- Ülekandmine
- Vastuvõtmine



Erinevused ettevalmistamisel

- Kui dokumendid on digitaalsed, on võimalik nende esialgne automaatne “kokku lugemine” ja “korrastamine”
- Kui kirjeldused on digitaalsed, on võimalik nende automaatne taaskasutamine arhiivikirjelduse loomisel
- Järeldus: digitaalsete kirjelduste olemasolul on võimalik laiendada kirjelduste mahtu ja detailsust
- Tehnilised tööd: dokumentide ja kirjelduste eraldamine EDHSist



Korrastamine ja kirjeldamine

- Kirjeldamisel toimub automaatne korrastamine – nimistu koostamisega paralleelselt toimub taustal failide tõstmine õigetes kaustadesse
- Kirjeldatakse ka dokumente – vajalikud on reeglid selle teostamiseks!
- Lisanduv vajadus on tehniliste metaandmete loomine ja kirjeldamine, võib-olla ka failide töötlemine
- Enne dokumentide edastamist arhiivi on võimalik automaatselt kontrollida osade nõuete täidetust juba asutuses

Ülekandmine

- Testülekande põhimõtte lisandumine – arhiivi edastatakse esmalt väike hulk dokumente ning arhiiviskeem, mida arhivaar saab põhjalikumalt analüüsida.
- Ülekanne toimub andmesidekanaleid kasutades, samamoodi tagasiside arhiiviskeemile ja dokumentidele ning nende kirjelduse kvaliteedile
- Järeldus: tagasiside kiirem

Automaatsed kontrollid

- Ainese täielikkus – kirjeldatud dokumentide hulk ja failide arv
- Ülekande täielikkus – kas välja saadetud ja vastu võetud failid on identsed
- Metaandmete täielikkus – kas kõik vajalikud elemendid on kirjeldatud
- Kirjelduste kvaliteet – kontrolli on võimalik osaliselt automatiseerida kindlate reeglite olemasolul
- Tehnilised kontrollid – failivormingud, tehnilised metaandmed, ...



Arutelu

- **Dokumentide ettevalmistamine pikaajaliseks säilitamiseks**
 - Millised tegevused dokumentide ettevalmistamisel nõuavad hetkel kõige rohkem manuaalset ja rutiinset tööd?
 - Milliste tegevuste automatiseerimine võiks kaasa tuua suurima kasu protsessi kiirendamiseks (andmete kirjeldamine, andmete eksport EDHSist, nimistu koostamine, infopakettide moodustamine ülekandmiseks jne)?
 - Millised tegevused on eeldatavasti liialt tehnilised asutuse arhivaari jaoks?
 - Kui suur võiks Teie hinnangul olla andmemah, mis ühe korruga pikaajaliseks säilitamiseks ette valmistatakse? Kui tihti seda tehakse?

- **Dokumentide ülekandmine arhiivi**
 - Milliste ülekandmist ettevalmistavate tegevuste (hindamine, sobiva ülekandemeetodi valimine, ülekandmise tähtaja määramine jne.) automatiseerimisest võiks asutusel enim kasu olla?
 - Kas mõne tegevuse automatiseerimine on Teie arvates mõeldamatu? Milliste?
 - Kas Teie jaoks oleks kasulik kasutada teiste asutuste kogemusi dokumentide ülekandmisel? Millist osa sellest?

- **Organisatsioonilised küsimused**
 - Milliste protsesside automatiseerimisel on kõige suurem oht kaotada intellektuaalne kontroll dokumentide ja nende kirjelduste üle?
 - Milliste protsesside automatiseerimine võiks enim ohustada arhiveeritavate andmete kvaliteeti ja autentsust?
 - Oletame, et Teil on kasutusel agentvahend, mis suudab suhelda teiste samalaadsete vahenditega teistes asutustes. Millised asutused (või asutuste grupid) need võiksid olla?

- **Sünergia teiste süsteemidega**

- Kas Te kasutate arhiveerimise teostamisel väliseid teadmusbasse?
- Millise sisuga teadmusbasidest võiks Teie töös kõige rohkem kasu olla?
- Milliste süsteemidega peaksid arhiveerimist hõlbustavad vahendid olema võimalised suhtlema?
 - Teie asutuses?
 - Teistes asutustes või riigis üldiselt?